Journal of Pure and Applied Algebra 92 (1994) 227–240
North-Holland

Rational Samuelson maps are univalent

L.A. Campbell

The Aerospace Corporation, 2350 East El Segundo Blvd, El Segundo, CA 90245, USA

Communicated by C.A. Weibel
Received 21 September 1992
Revised 6 March 1993

Abstract

Campbell, L.A., Rational Samuelson maps are univalent, Journal of Pure and Applied Algebra 92 (1994) 227–240.

A differentiable self-mapping of n-space is Samuelson if the leading principal minors of its Jacobian matrix vanish nowhere. The principal result of this paper is that a continuously differentiable Samuelson map of real n-space to itself, with component functions that have algebraic graphs, is bijective and decomposable into n semialgebraic diffeomorphisms, each of which changes only a single different coordinate. In particular, everywhere defined real rational Samuelson maps are univalent.

1. Introduction

A *Samuelson map* is a map, given by n functions of n variables, with the property that the Jacobian matrix has leading principal minors that vanish nowhere. That is, if $F = (f_1, \ldots, f_n)$, then F is Samuelson if the determinants $\mu_r = \det(\partial f_i/\partial x_j)_{i,j=1,r}$ are everywhere nonzero, for $r = 1, \ldots, n$. The definition of a Samuelson map is clearly coordinate-system dependent. The name arises from a 1953 paper on the theory of general equilibrium in economics by Paul A. Samuelson, the noted economist and Nobel laureate [20]. In it, he suggested that the above condition on minors for a differentiable self-map of $\mathbb{R}^n$ would guarantee univalence. The term "Samuelson map" is applicable in a variety of contexts, for instance to maps that are defined only on subsets of n-space, to holomorphic maps and complex partial derivates, or to polynomial maps with coefficients in a commutative ring [7].

In 1965, Gale and Nikaidô published a simple example of a Samuelson map of $\mathbb{R}^2$ to itself which is not univalent: $F(x, y) = (e^{2x} - y^2 + 3, 4e^{2x}y - y^3)$ [9].

Correspondence to: L.A. Campbell, M1-102 Aerospace, P.O. Box 92957, Los Angeles, CA 90009, USA.

SSDI 0022-4049(93)E0083-G

Example 1.1. Consider the slightly more general map $F = (f, g)$ with the component functions $f(x, y) = h(x) - y^2 + 3$ and $g(x, y) = 4yh(x) - y^3$. The Jacobian matrix of this map is

$$\begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix} = \begin{bmatrix} h'(x) & -2y \\ 4yh'(x) & 4h(x) - 3y^2 \end{bmatrix}$$

and hence $\mu_1 = h'(x)$ and $\mu_2 = 4h(x)h'(x) + 5y^2h'(x)$. This F is Samuelson if $h'(x)$ does not vanish and $h(x) > 0$, both clearly true if $h(x) = e^{2x}$. The only other property of e^{2x} that h needs to share in order to produce an easy counterexample to univalence is $h(0) = 1$. For then $F(0, y) = (4 - y^2,\ 4y - y^3)$ and $F(0, 2) = (0, 0) = F(0, -2)$. In particular, one can choose $h(x) = x + \sqrt{1 + x^2}$ to produce an example of a Nash (real analytic and semialgebraic) Samuelson map, defined on all of $\mathbb{R}^2$, that is not univalent.

Note, however, that one cannot use the same trick to produce a similar rational example, since no rational function $h(x)$ can satisfy both $h(x) > 0$ and $h'(x) \neq 0$ for all real x. Thus Example 1.1 suggests the possibility that *rational* Samuelson maps, defined on all of $\mathbb{R}^n$, must be univalent (one-to-one). The main theorem of this paper states that this is true, and even asserts a somewhat stronger statement.

Main Theorem. *Let $F = (f_1, \ldots, f_n)$ be a map defined on all of $\mathbb{R}^n$, with values in $\mathbb{R}^n$, and with continuous first-order partial derivatives ($\mathscr{C}^1$). If F is Samuelson and each component function f_i has an algebraic graph, then F is bijective (one-to-one and onto).*

Here the graph of f_i is $\{(x_1, \ldots, x_n, z) \in \mathbb{R}^{n+1} \mid z = f_i(x_1, \ldots, x_n)\}$, and the condition is that this must be an algebraic set (the zero-set of a collection of polynomials) for each i. If $f_i = p/q$, where p and q are polynomials with q vanishing nowhere, then the graph of f_i is the algebraic set where $zq = p$. So everywhere defined rational Samuelson maps, and thus polynomial Samuelson maps, are bijective. To see that the theorem does *not* apply to the map of Example 1.1, with $h(x) = x + \sqrt{1 + x^2}$, observe that if the graph of $h - y^2 + 3$ were algebraic, then so would be the graph of h. But the graph of h is the set $(z - x)^2 = 1 + x^2, z > 0$, which is only one connected component of the (irreducible) algebraic curve $z^2 - 2xz - 1 = 0$, and so h is only semialgebraic (its graph is defined by polynomial equalities and inequalities).

Some remarks on terminology are in order. The terms "univalent", "one-to-one", and "injective" are synonyms, as are "onto" and "surjective". Choices from among synonyms are made on purely stylistic grounds. "Everywhere defined rational map" means a map whose components can be written as quotients of polynomials, with the denominators vanishing nowhere in the domain of definition; such maps have elsewhere been called "morphisms" [5] and "regular maps" [6]. Samuelson maps have also been called "maps satisfying the Samuelson condition" [15], and "NVL-transformations" [21].

The statement of the main theorem given above is amplified later. The full statement asserts that if F satisfies the given conditions, then F can be factorized as the composition $F = V_n \circ \cdots \circ V_1$ of semialgebraic $\mathscr{C}^1$ diffeomorphisms V_i, where V_i fixes all the coordinates except x_i. If, in addition, F is of class $\mathscr{C}^k$ ($k > 1$ or $k = \infty$), then each V_i, and its inverse, is of class $\mathscr{C}^k$. Since semialgebraic $\mathscr{C}^\infty$ maps are real analytic [6, Proposition 8.1.7], it is also true that F real analytic implies that the V_i are real analytic.

The following sections of this paper consider prior related results, versions of the main theorem over other fields than $\mathbb{R}$, automatic surjectivity results for injective algebraic maps, maps whose components have algebraic graphs, particulars of the case $n = 2$, the proof for general n, and some loose ends.

2. Related work

The paper [9], in which Gale and Nikaidô gave their example of a non-univalent Samuelson map, is actually devoted primarily to the proof that a map is one-to-one if its Jacobian matrix is everywhere a P-matrix (one for which *every* principal minor is *positive*). This result holds for general differentiable maps defined on a rectangular region $\Omega \subseteq \mathbb{R}^n$. This is the prototype of the *global univalence theorems* that are the subject of Parthasarathy's book [17]. The book deals extensively with similar results, in which univalence is a consequence of pointwise conditions on the Jacobian matrix, growth conditions on its norm, and/or conditions on the map or its Jacobian matrix at the boundary points of a region. Recent work in this area is exemplified by [16].

The *Jacobian Conjecture* asserts that a map $F: \mathbb{C}^n \to \mathbb{C}^n$ which is polynomial in the complex coordinates of $\mathbb{C}^n$ is one-to-one and onto (in which case it has a polynomial inverse), if, and only if, its Jacobian matrix has a non-zero *constant* determinant. Good surveys of this longstanding problem (unsolved even for $n = 2$) can be found in [3,8,19,23]. The *Real Jacobian Conjecture* asserts that if a polynomial (or more generally, everywhere defined rational) map $F: \mathbb{R}^n \to \mathbb{R}^n$ has a Jacobian matrix whose determinant vanishes nowhere, then it is one-to-one and onto (with an inverse that is not necessarily polynomial). Connections have been drawn to the theory of ordinary differential equations [11–13]. The problem remains open even for $n = 2$, and even if the determinant is assumed to be constant.

Real planar ($n = 2$) maps with non-vanishing Jacobian determinant have been the subject of numerous articles (e.g., [4, 11]). For $\mathscr{C}^1$ maps, if the Jacobian matrix J has a non-vanishing determinant and one can choose a single entry from each column of J that does not vanish anywhere, than the map is injective [9, 14]. For polynomial maps, the conditions $\det(J) > 0$ and $\operatorname{trace}(J)$ nowhere zero imply injectivity [13]. The following example, although not Samuelson, is particularly relevant, since it is rational except at one point.

Example 2.1 (Meisters [11]). Consider the map $F = (f, g)$, where

$$f(x, y) = (10x^3y^2 - x^5 - 5xy^4)/\sqrt{5}(x^2 + y^2)^2,$$
$$g(x, y) = (10x^2y^3 - y^5 - 5x^4y)/\sqrt{5}(x^2 + y^2)^2.$$

F is defined everywhere except at $(0, 0)$, and its Jacobian determinant is identically 1. Define $F(0, 0) = (0, 0)$. Then the extended map F has x- and y-partials at the origin, and defines a (non-$\mathscr{C}^1$) map from $\mathbb{R}^2$ to $\mathbb{R}^2$, for which $\det(J)$ exists everywhere, and is 1 except at $(0, 0)$, where it is $1/5$. However, the map is *not* univalent ($F(1, 0) = F(\cos 2\pi/5, \sin 2\pi/5)$).

For $n = 3$ a matrix with positive determinant and all positive entries need not be a P-matrix. This example is taken, attribution and all, from [14].

Example 2.2 (Ravindran [18]). Define the map $F : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$F(x,y,z) = (e^{2x} - y^2 + 3, 4e^{2x}y - y^3, 4(10 + e^{2x})\cosh(y)\sinh(100z)).$$

Let

$$A = \frac{1}{3}\begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Treat vectors in $\mathbb{R}^3$ as column vectors, and put $G(x) = A^{-1}F(Ax)$. Then the Jacobian matrix J of G has a positive determinant, and all the entries of J are positive, but of course G is not univalent ($F(0, 2, 0) = F(0, -2, 0)$).

A number of results that will be used come under the general heading of *real algebraic geometry* [6]. Injective, everywhere defined rational maps of $\mathbb{R}^n$ to itself are automatically surjective [5], as are injective, everywhere defined rational maps of a real algebraic variety to itself [6], as are injective, continuous maps of $\mathbb{R}^n$ to itself with algebraic graphs [10]. This last result figures significantly in the proof of the main theorem.

Complex polynomial Samuelson maps are of a particularly simple form. These are maps from $\mathbb{C}^n$ to itself that are polynomial in the complex coordinates and whose Jacobian matrix, computed using complex partials, has nowhere vanishing principal minors. Since the minors are polynomials in the complex coordinates and vanish nowhere, they are necessarily constant. Such maps have, in fact, been classified by van den Essen and Parthasarathy [24], who showed that they are not only bijective but also that they can be written as the composition of a diagonal linear map and n elementary maps (maps with Jacobian determinant 1 and only a single coordinate that changes). The proof is algebraic, and applies more generally to polynomial maps (with constant leading principal minors) defined over suitable commutative rings. The present author extended and generalized these decomposability results to classes of non-polynomial maps with minors that need not be constant [7]. The idea of decomposing a map into basic maps (invertible maps with only a single coordinate changing) is central to the proof of the main theorem.

3. Other fields

This section addresses the generalization of the main theorem to other fields, that is, to maps $K^n \to K^n$ where K is a field other than $\mathbb{R}$. Throughout this section it is assumed that the theorem is known for $\mathbb{R}$. The conditions $\mathscr{C}^1$ and Samuelson, with components whose graphs are algebraic, are perfectly meaningful over any real closed field [6]. Let $\mathscr{L}$ be the first order language of ordered fields. A consequence of the *Tarski–Seidenberg principle* [6, 22] is that the theory of real closed fields is complete; that is, any sentence of $\mathscr{L}$ is true in one real closed field if, and only if, it is true in every real closed field.

Theorem 3.1. *The main theorem is true over any real closed field.*

Proof. This proof will only be a sketch. The main theorem cannot directly be expressed in $\mathscr{L}$. For instance, it is supposed to be true for all n, but the language does not allow quantification over the integers (only over field elements). Also, maps and functions are not first order objects. However, given any fixed positive integer m one can speak of the existence of m objects by repeating existential quantifiers ($\exists a_1 \exists a_2 \ldots \exists a_m$). And one can make statements about algebraic sets by expressing them in terms of statements about the coefficients of the polynomials defining the algebraic set, but of course only if one has an *a priori* bound on the number of polynomials and their degrees. The idea of the proof is to create sentences $\phi[n, m, d]$, parameterised by positive integers, whose meaning is that of the following (English) sentence: For any n algebraic sets, each the zero-set of at most m polynomials in $n + 1$ variables of degree at most d, such that the algebraic sets are the graphs of functions, and for which the map with those functions as components is $\mathscr{C}^1$ and Samuelson, that map is injective. The sentences $\phi[n, m, d]$ of course do not have variables corresponding to m, n, and d; rather m, n, and d determine how many variables, clauses, etc. appear in the formulas. The main theorem is equivalent to the simultaneous truth of all the sentences $\phi[n, m, d]$. Since the theorem is true over $\mathbb{R}$, each sentence is true, by completeness, over any real closed field K, and hence they are all true over K, which yields the theorem for K. $\square$

A look at the above argument will also reveal that is can be extended to prove decomposability (see Section 7) as well. (More precisely, an extension of the argument yields the $\mathscr{S}^k$ case for finite k, and the $\mathscr{S}^\infty$ case follows trivially. Over a general real closed field, Nash functions are defined as the $\mathscr{S}^\infty$ functions, obviating any questions about analyticity. See Section 7 for the appropriate definitions.)

Over the complex numbers $\mathbb{C}$, one would like a generalization in terms of holomorphic functions and complex partial derivatives. However, requiring the graphs of the components to be algebraic rules out all but rational maps. Furthermore, any everywhere defined rational map is polynomial, because the nowhere vanishing denominators must be non-zero constants. As mentioned, van den Essen and Parthasarathy completely classified polynomial Samuelson maps over $\mathbb{C}$ [24]. Their proof applies immediately to any algebraically closed field of characteristic zero (alternatively, one could use the *Lefshetz* principle to transfer the result from $\mathbb{C}$ to an arbitrary algebraically closed field of characteristic zero). We record this as the following proposition.

Proposition 3.2 (van den Essen and Parthasarathy [24]). *The main theorem is true for everywhere defined rational maps over any algebraically closed field of characteristic zero. Such maps are necessarily polynomial.* □

Over any field one can ask whether rational Samuelson maps are injective. However, it is clear that this is not the suitable formulation of the question, even for real fields such as the rationals $\mathbb{Q}$.

Example 3.3. The map $f: \mathbb{Q} \to \mathbb{Q}$ given by the function $f(x) = x^3 - 6x$ is a polynomial Samuelson map over $\mathbb{Q}$ which is not injective. It is Samuelson because the zeros of $f'(x)$, namely $\pm\sqrt{2}$, do not belong to $\mathbb{Q}$.

Of course the main problem with f in the above example is that f' changes sign, and hence f is not monotone (over either $\mathbb{Q}$ or $\mathbb{R}$). One relevant result is that for any field K, any polynomial map $f: K^n \to K^n$ with a Jacobian determinant that vanishes nowhere on K^n, and any $y \in K^n$, the number of points in the fiber $f^{-1}(y)$ is bounded by the degree of the field extension $K(x_1, \ldots, x_n): K(f_1, \ldots, f_n)$ [1, 2].

4. Automatic surjectivity

There are a number of circumstances in which the fact that a map of some mathematical entity to itself is one-to-one automatically implies that it is onto. Examples are self-maps of finite sets, linear self-maps of finite-dimensional vector spaces, and endomorphisms of affine algebraic varieties over an algebraically closed field. For the proof of the main theorem, "automatic surjectivity" results of this type are required for real non-linear maps that are, in some sense, algebraic.

In [5], Białynicki-Birula and Rosenlicht showed that one-to-one everywhere defined rational map of $\mathbb{R}^n$ to itself is onto. In [6, Theorem 11.4.2], this was extended to everywhere defined rational maps of an irreducible real algebraic set to itself. The strongest results were obtained in [10], but only for maps of $\mathbb{R}^n$ to itself. Kurdyka and Rusek showed that a continuous semialgebraic map of $\mathbb{R}^n$ to itself is one-to-one, provided that its graph behaves appropriately at infinity. Specifically, the condition is that there exist an algebraic compactification X of $\mathbb{R}^n \times \mathbb{R}^n$, and an arcwise symmetric set $E \subset X$, whose finite portion $(E \cap \mathbb{R}^{2n})$ is the graph of the map. Here an arcwise symmetric subset $E \subset X$ is one that contains all of an analytic arc $\gamma:(-1, 1) \to X$ if it contains half of it $(\gamma(-1, 0)) \subset E)$. The complete strength of this result will not be used, although it is possible that one could use it to extend the results of this paper. Because algebraic sets are arcwise symmetric, one consequence of this result is automatic surjectivity for any continuous injective map whose graph is either an algebraic set or the only unbounded component of an algebraic set. Ignoring the complication of compact components, this yields [10, Corollary 4.2], recorded here as a proposition for later use.

Proposition 4.1 (Kurdyka and Rusek [10]). *Let $F:\mathbb{R}^n \to \mathbb{R}^n$ be an injective, continuous mapping with algebraic graph. Then F is surjective.* $\square$

5. Component-wise algebraic maps

Denote by $\mathscr{A}(\mathbb{R}^n)$, or $\mathscr{A}$ for short, the set of real functions of n real variables with algebraic graphs. A function f belongs to $\mathscr{A}$ precisely when its graph $\Gamma = \bar{\Gamma}$, the Zariski closure of Γ (that, is the smallest algebraic set containing Γ). Another, useful, way of putting this is that $f \in \mathscr{A}$ if for any $x \in \mathbb{R}^n$ there is a unique point $(x, y) \in \bar{\Gamma}$.

Example 5.1. Let $n = 1$ and $y = x^{(3k+1)/3}$, for k a non-negative integer. Then y is precisely $\mathscr{C}^k$ (continuously differentiable of order k). This is true even for $k = 0$, with y merely continuous in that case. Furthermore, the graph of y is the algebraic set $y^3 = x^{3k+1}$, so $y \in \mathscr{A}$.

Example 5.2. Let $n = 1$ and define $y = f(x)$ by $f(x) = 1/x$ for $x \neq 0$ and $f(0) = 0$. Then the graph of f is the union of the hyperbola $xy = 1$ and the point $(0, 0)$ and hence is algebraic (but reducible, defined, for instance, by $(xy - 1)(x^2 + y^2) = 0$). Note that f is not continuous.

Example 5.3. Functions of a single variable that are not in $\mathscr{A}$ include many semialgebraic functions, such as $y = \sqrt{1 + x^2}$, whose graph is only one connected component of its Zariski closure $y^2 = 1 + x^2$. Note that y is Nash (real analytic as well as semialgebraic).

Functions with algebraic graphs are not well behaved under the usual operations on real functions, such as the arithmetic operations or composition of functions. The basic reason for this is that the projection of an algebraic set (from, say $\mathbb{R}^n$ to $\mathbb{R}^m$, with

$m < n$) is not in general algebraic. In contrast, the projection of a semialgebraic set is semialgebraic (and one obtains algebras of semialgebraic functions, and composition operators [6]).

Example 5.4. Let $v = x^{1/3}$ and $y = f(v) = v + v^2 = x^{1/3} + x^{2/3}$. Then y, considered as a function of x, does not have an algebraic graph. For

$$y^3 = (v + v^2)^3 = v^3(1 + v)^3 = x(1 + 3(v + v^2) + v^3) = x(1 + 3y + x).$$

Thus y satisfies the equation $P(y, x) = 0$, where P is the (irreducible) polynomial $y^3 - 3yx - (x + x^2)$. $P = 0$ is the equation of the Zariski closure of the graph of y, but it contains the points $(x = 1, y = 2)$ and $(x = 1, y = -1)$ so it is not equal to that graph. On the other hand, v, v^2 and $1 + v$ are in $\mathscr{A}$, and $f(v)$ is a polynomial in v. But $y = f(v) = v + v^2 = v(1 + v)$, which shows that $\mathscr{A}$ is not closed under composition, addition, or multiplication.

Maps each of whose components belong to $\mathscr{A}$ will be called *component-wise algebraic maps* (abbreviated CAM). Thus if $F:\mathbb{R}^n \to \mathbb{R}^m$ and $F = (f_1, \ldots, f_m)$, then F is a CAM if each $f_i \in \mathscr{A}(\mathbb{R}^n)$. This definition, like that of a Samuelson map, is coordinate system dependent, but is independent of the order of enumeration of the coordinates. It is more stringent than the (coordinate free) requirement that graph of F ($\Gamma(F) \subset \mathbb{R}^n \times \mathbb{R}^m$) be algebraic.

Example 5.5. Let $m = n = 2$ and define a map from $\mathbb{R}^2$ to $\mathbb{R}^2$ by $F(x,y) = (x^3, y - x - x^2)$. F is a polynomial map, hence it is component-wise algebraic, and thus also has an algebraic graph. Introduce coordinates (u, v) in the codomain. The equations $u = x^3$, $v = y - x - x^2$ can be solved, and the solution written as $x = u^{1/3}$, $y = v + u^{1/3} + u^{2/3}$. Thus F has an inverse G. Since the graph of G is the same as that of F (upon identifying domain and codomain properly), G has an algebraic graph. However, the second component of G, namely $v + u^{1/3} + u^{2/3}$ does not belong to $\mathscr{A}$, since Example 5.4 shows that $u^{1/3} + u^{2/3} \notin \mathscr{A}$. This shows that if G has an algebraic graph it is not necessarily a CAM, and that if F is a CAM, F^{-1} is not necessarily a CAM.

The reason for considering component-wise algebraic maps is that they have properties not shared by maps that just have an algebraic graph. The property of interest is that if F is a CAM, then so is the map obtained by dropping any subset of the components.

6. The planar case

Although it contains all the essential elements of the proof for general n, a proof of the main theorem for $\mathscr{C}^1$ maps in the planar case $n = 2$ involves fewer details, and is thus easier to follow. After the presentation of the proof, consequences particular to the planar case are developed. The notation $\mathcal{S}^1$ denotes $\mathcal{C}^1$ maps that are semialgebraic as well, and an $\mathcal{S}^1$ diffeomorphism is an invertible $\mathcal{S}^1$ map whose inverse is also $\mathcal{S}^1$.

Theorem 6.1. *A $\mathcal{C}^1$ Samuelson component-wise algebraic map $F:\mathbb{R}^2 \to \mathbb{R}^2$ is bijective and can be written as a composition $F = U \circ V$, where U fixes the first coordinate, V fixes the second, and both U and V are $\mathcal{S}^1$ diffeomorphisms.*

Proof. Let F such a map. Write $F = (f, g)$, where $f(x, y)$ and $g(x, y)$ are functions with algebraic graphs. Attempt to factor the map as $F = U \circ V$, where $V = (f(x, y), y)$ and $U = (x, h(x, y))$. Use subscripts to denote partials. Since F is Samuelson, $\mu_1 = f_x$ and $\mu_2 = f_x g_y - f_y g_x$ vanish nowhere. Since f_x is either always positive or always negative, f is monotone in x for each fixed y, and thus V is injective. Since F is a CAM, V has an algebraic graph, so by Proposition 4.1 it is bijective. Let $V^{-1} = (\theta(x, y), y)$. Since V is semialgebraic and has a nowhere vanishing Jacobian determinant, the $\mathcal{S}^1$ inverse function theorem shows that V is an $\mathcal{S}^1$ diffeomorphism [6, Proposition 2.9.5]. By definition θ satisfies $f(\theta(x, y), y) = x$. In order to have $F = U \circ V$, it must be the case that $g = h \circ V$; that is, $g(x, y) = h(f(x, y), y)$. Substitute $\theta(x, y)$ for x to obtain $h(x, y) = g(\theta(x, y), y)$. This defines h in the only possible way, and it is clear that with that definition, $F = U \circ V$.

Next, observe that $\mu_2 = (\mu_1)(h_y \circ V)$. For $g = h \circ V$, so $g_x = h_x(f, y) f_x$ and $g_y = h_x(f, y) f_y + h_y(f, y)$ and thus $\mu_2 = f_x g_y - f_y g_x = f_x h_y(f, y)$ as claimed. It follows that $h_y = (\mu_2/\mu_1) \circ V^{-1}$ vanishes nowhere, which implies that U is injective. But then $F = U \circ V$ is a continuous injective map of $\mathbb{R}^2$ to itself with algebraic graph and hence it is bijective by Proposition 4.1. This implies that U itself is bijective. Since U semialgebraic (because h is a composition of semialgebraic functions) and has nowhere vanishing Jacobian determinant (because h_y is nowhere zero), it follows from the $\mathcal{S}^1$ inverse function theorem that U is an $\mathcal{S}^1$ diffeomorphism. $\square$

Corollary 6.2. *Let $F:\mathbb{R}^2 \to \mathbb{R}^2$ be a $\mathcal{C}^1$ component-wise algebraic map with nowhere vanishing Jacobian determinant. If any one of the four entries of its Jacobian matrix vanishes nowhere, then F is bijective.*

Proof. Let $F(x, y) = (f(x, y), g(x, y))$. Interchange the coordinates x, y, or the components f, g, or both, to bring the nowhere vanishing entry of the Jacobian into the $(1, 1)$ position. The resulting map satisfies the conditions of Theorem 6.1, hence is bijective, and thus so is F. $\square$

Corollary 6.3. *Let $F:\mathbb{R}^2 \to \mathbb{R}^2$ be an everywhere defined rational map, with nowhere vanishing Jacobian determinant, and suppose that for some vectors $u, v \in \mathbb{R}^2$ the function $u^t Jv$ vanishes nowhere. Then F is bijective.*

Table 1
Injectivity for $F : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ with det (J) nowhere zero

F	Ω	Condition for injectivity	Reference
$\mathscr{C}^1$	OR	$f_x g_y$ nowhere zero	[9, Theorem 7]
$\mathscr{C}^1$	OR	f_x and g_y one-signed	[9, Theorem 7]
$\mathscr{D}$	OR	$\exists u \neq \in \mathbb{R}^2$, with $\det(J)$, $u^t J$ one-signed	[9, Theorem 8]
$\mathscr{C}^1$	$\mathbb{R}^2$	one of $f_x f_y, f_x g_y, g_x f_y, g_x g_y$ nowhere zero	[14, Corollary 1]
$\mathscr{P}$	$\mathbb{R}^2$	$\det(J) > 0$ and trace(J) nowhere zero	[13, p. 379]
CAM^1	$\mathbb{R}^2$	some entry in J nowhere zero	Collary 6.2
$\mathscr{R}$	$\mathbb{R}^2$	$\exists u, v \in \mathbb{R}^2$, with $u^t J v$ nowhere zero	Corollary 6.3

Proof. Let $F = (f, g)$. Obviously, neither u nor v can be the zero vector. Let A and B be invertible matrices, with u^t the first row of A, and v first column of B. The map $G = AF(Bx)$, where juxtaposition indicates matrix multiplication, has Jacobian $G' = A(F' \circ \beta)B = (AF'B) \circ \beta$, where β is the linear map $x \rightsquigarrow Bx$. G is rational, has nowhere vanishing Jacobian determinant, and its upper left element is the nowhere vanishing function $(u^t J v) \circ \beta$. By the previous corollary, G is bijective, hence so is F. Note that his argument does not work if F is only $\mathscr{C}^1$ and a CAM, because G need not be a CAM. □

Table 1 is intended to place the above corollaries in context. It summarizes them along with some other cases in which a map F from an open set $\Omega \subseteq \mathbb{R}^2$ to $\mathbb{R}^2$ is known to be injective. Assume that $F = (f, g)$, that F has partial derivatives throughout Ω, and that the Jacobian matrix, J, of F has a determinant that vanishes nowhere in Ω. Additional restrictions are imposed on F and/or Ω in the different cases. $\mathscr{D}$ denotes differentiable maps (Fréchet derivative), CAM^1 component-wise algebraic maps that are $\mathscr{C}^1$, $\mathscr{R}$ everwhere defined rational functions, and $\mathscr{P}$ is the class of polynomial maps. Note that Example 2.1 is not in any of the classes. Ω is either an open rectangle (OR) with sides parallel to the parallel to the axes (and possibly equal to $\mathscr{R}^2$), or all of $\mathbb{R}^2$. The term "one-signed" is used for a function that is everywhere non-negative or everywhere non-positive, and for a vector of functions with one-signed components.

The table displays only some selected, and partially overlapping, results. The attempt is to show the known results that are closest in their form to that of the corollaries, that is, involve simple sign conditions on the entries of J. A noticeable difference between earlier results and the present corollaries is that the corollaries involve one, rather than two, sign conditions on the entries of J (counting a sign condition on a product as two sign conditions).

7. The general case

A map $F: \mathbb{R}^n \to \mathbb{R}^n$ is *decomposable* if it can be written as a composition $F = V_n \circ \cdots \circ V_1$, where each map V_i is invertible and changes only the ith coordinate

[7]. Such a decomposition is unique, when it exists [7]. Let $\mathscr{S}^k$ denote the class of $\mathscr{C}^k$ semialgebraic maps. If the factors V_i and their inverses are $\mathscr{S}^k$ ($k > 0$ or $k = \infty$), the map is *$\mathscr{S}^k$-decomposable.* These definitions are coordinate system dependent, and even depend on the order in which the coordinates are enumerated. Clearly, a decomposable map is bijective.

Lemma 7.1. *Suppose that F is $\mathscr{S}^1$-decomposable and that F is also of class $\mathscr{C}^k$ ($k > 0$ or $k = \infty$). Then the factors in the $\mathscr{S}^1$ decomposition of F, and their inverses, are of class $\mathscr{S}^k$, and thus F is $\mathscr{S}^k$-decomposable.*

Proof. The proof is by induction on the length of a decompsition $F = V_n \circ \cdots \circ V_m$, where V_i changes only the ith coordinate and V_i and its inverse are of class $\mathscr{S}^1$. Observe that $V_m = (x_1, \ldots, x_{m-1}, f_m, x_{m+1}, \ldots, x_n)$. Since F is of class $\mathscr{C}^k$, so is f_m. But then V_m is a $\mathscr{C}^k$ map which is a $\mathscr{C}^1$ diffeomorphism. By the inverse function theorem, its inverse is also of class $\mathscr{C}^k$. Also, the inverse of any semialgebraic map is semialgebraic. If $m = n$ this completes the proof. Otherwise, consider the map $G = F \circ V_m^{-1}$. It is $\mathscr{C}^k$ and has the $\mathscr{S}^1$ decomposition $V_n \circ \cdots \circ V_{m+1}$. By induction, the factors of that decomposition, and their inverses, are also $\mathscr{S}^k$. $\square$

The above proof can, in fact, be mimicked for any class of maps that is defined component-wise, is closed under composition, and admits an inverse function theorem. In particular, one could prove the analogous result for real analytic maps, but $\mathscr{S}^\infty$ is the same as Nash (real analytic and semialgebraic) over the reals [6, Proposition 8.1.7].

Theorem 7.2 (Main Theorem). *A $\mathscr{C}^k$ ($k > 0$ or $k = \infty$) Samuelson component-wise algebraic map from $\mathbb{R}^n$ to $\mathbb{R}^n$ is bijective and $\mathscr{S}^k$-decomposable.*

Proof. By Lemma 7.1, it is enough to consider only the $\mathscr{C}^1$ case. Let $F = (f_1, \ldots, f_n)$ be a map from $\mathbb{R}^n$ to $\mathbb{R}^n$ that satisfies the hypotheses. To prove the desired result requires factorizing F as a composition $F = V_n \circ \cdots \circ V_1$, where each V_i is an $\mathscr{S}^1$ diffeomorphism of $\mathbb{R}^n$ in which only the i-th coordinate changes. Clearly, $V_1 = (f_1, x_2, \ldots, x_n)$. As in the planar case, V_1 is an $\mathscr{S}^1$ diffeomorphism, with inverse $(\theta(x_1, \ldots, x_n), x_2, \ldots, x_n)$. Put $U_1 = F \circ V_1^{-1}$, so that $F = U_1 \circ V_1$. It is easy to see that U_1 is $\mathscr{S}^1$ and has first component x_1; that is, it fixes the first coordinate. Suppose, by induction, that at a certain stage in the factorization $F = U_k \circ V_k \circ \cdots \circ V_1$, where the V_i are as specified for $i \leq k$ and U_k is $\mathscr{S}^1$ and fixes the first k coordinates. Write $U_k = (x_1, \ldots, x_k, h_k, \ldots)$; in other words, let h_k be component number $k+1$ of U_k. Assume for the moment that it has been established that the partial of h_k with respect to x_{k+1} vanishes nowhere. Fix the coordinates $x_{k+2}, \ldots, x_n$ at a value $a = (a_{k+2}, \ldots, a_n)$ and drop all but the first $k+1$ components of F. Let F_a^k denote the resulting map from $\mathbb{R}^{k+1}$ to itself. Make similar definitions for U_k and V_i (for $i \leq k$). Observe that $F_a^k = U_{k,a}^k \circ V_{k,a}^k \circ \cdots \circ V_{1,a}^k$, because the V_i involved do not change the coordinates of a. The $V_{i,a}^k$ are $\mathscr{S}^1$ diffeomorphisms of $\mathbb{R}^{k+1}$. Since

$U^k_{k,a} = (x_1, \ldots, x_k, h(x_1, \ldots, x_{k+1}, a_{k+2}, \ldots, a_n))$, and the partial of h_k with respect to x_{k+1} vanishes nowhere, $U^k_{k,a}$ is injective, and hence so is F^k_a. Since F is a CAM, so is F^k_a, and thus F^k_a is bijective, by Proposition 4.1. But then $U^k_{k,a}$ is also bijective (the critical point being that it is onto). This means that there is an inverse for h_k with respect to the x_{k+1} coordinate, that is, a function $\varphi(x_1, \ldots, x_{k+1})$ such that $h_k(x_1, \ldots, x_k, \varphi) = x_{k+1}$. Recall that the coordinates $x_{k+2}, \ldots, x_n$ have been fixed at the value a. Extend φ to a function of all n variables by allowing a to vary (that is, combine the functions φ for different values of a). The result is a solution to the functional equation $h_k(x_1, \ldots, x_k, \varphi, x_{k+2}, \ldots, x_n) = x_{k+1}$. By the $\mathscr{S}^1$ implicit function theorem [6], φ is $\mathscr{S}^1$ in all its variables. Define $V_{k+1} = (x_1, \ldots, x_k, \varphi, x_{k+2}, \ldots, x_n)$ and $U_{k+1} = U_k \circ V_{k+1}^{-1}$. This provides the factorization $F = U_{k+1} \circ V_{k+1} \circ \cdots \circ V_1$ for the next induction step, as U_{k+1} is $\mathscr{S}^1$ and preserves the first $k+1$ coordinates by construction. At the last step $k+1 = n$ and U_{k+1} is the identity since it preserves all the coordinates. This yields the desired factorization $F = V_n \circ \cdots \circ V_1$. It remains to verify that $\partial h_k/\partial x_{k+1}$ vanishes nowhere. Let $W_k = V_k \circ \cdots \circ V_1$. Then W_k is $\mathscr{S}^1$ and for any fixed $a = (a_{k+2}, \ldots, a_n)$, one has $F^k_a = U^k_{k,a} \circ W^k_{k,a}$. By the chain rule $(F^k_a)' = ((U^k_{k,a})' \circ W^k_{k,a})(W^k_{k,a})'$. Taking determinants, $\mu_{k+1} = (\partial h_k/\partial x_{k+1}) \det(W^k_{k,a})'$. Since μ_{k+1} does not vanish anywhere, the same must be true of $\partial h_k/\partial x_{k+1}$. Since this is true for all values of a, the partial does not vanish anywhere when considered as a function of all n variables. □

Note that any $\mathscr{S}^k$ decomposition of F is the same as its $\mathscr{S}^1$ decomposition, because a decomposition of the kind considered here is unique—without regard to differentiability. Observe also that any real analytic F to which the theorem applies will above a Nash inverse and Nash factors.

8. Loose ends

The inverse of an invertible Samuelson map is itself Samuelson, provided one reverses the order of the coordinates [21]. However, the inverse of a component-wise algebraic Samuelson map is not necessarily component-wise algebraic. This is demonstrated by the following example (which supplements Example 5.5, which was not Samuelson and even had a singular Jacobian matrix along one axis). Note that the inverse of any polynomial Samuelson map is necessarily Nash, and so are the factors that occur in its decomposition.

Example 8.1. Let $F: \mathbb{R}^2 \to \mathbb{R}^2$ be the polynomial map $F = (f, g)$ with components $f(x, y) = x + x^3y^2$ and $g(x, y) = y + y^3x^2$. Then F is Samuelson (direct calculation shows $\mu_1 = 1 + 3x^2y^2$ and $\mu_2 = 1 + 6x^2y^2 + 5x^4y^4$). Let $F^{-1} = (p, q)$. Then $p + p^3q^2 = x$ and $q + q^3p^2 = y$. Hence $py = qx$. But then $x^3 = x^2(p + p^3q^2) = px^2 + p^5y^2$. This establishes $P(x, y, p) = px^2 + p^5y^2 - x^3$ as an obviously irreducible polynomial whose zero-set contains the graph of p. The algebraic set $P(x, y, p) = 0$ in

(x, y, p)-space turns out to be an "umbrella" [6]. That is, contains a two-dimensional "shade" plus a one-dimensional "handle"—the straight line $(0, 0, p)$. To see this, distinguish the case $(x, y) = (0, 0)$ and $(x, y) \neq (0, 0)$. The equation $P(0, 0, p) = 0$ is satisfied by any p, yielding the handle. For $x \neq 0$, the partial of P with respect to p is $x^2 + 5p^4y^2 > 0$, so P is monotone increasing and has at most one solution $P = 0$. Since it has one solution arising from the graph of p, it has exactly one solution. For $x = 0$ and $y \neq 0$, $P(0, y, p) = p^5y^2 = 0$ has exactly one solution, namely $p = 0$. These solutions, together with the point $(0, 0, 0)$ (which is also a point of the handle) constitute the shade of the umbrella, and the shade is exactly the graph of p. Thus the graph of p is *not* algebraic, because its Zariski closure contains the handle as well. So F^{-1} is not component-wise algebraic. (By symmetry, q also does not have an algebraic graph.)

In contrast, the example $F = (f, g)$, with $f(x, y) = (1 + 3y^2)(x + x^3)$ and $g(x, y) = (1 + 3x^2)(y + y^3)$, that was considered in [7], does turn out to have a component-wise algebraic inverse. (If $F^{-1} = (p, q)$, then the graph of p is the algebraic set satisfying $27y^2(p + p^3) = (x - p - p^3)(x + 2p + 2p^3)^2(1 + 3p^2)^2$.) Thus the main theorem applies to the inverse. This raises the following question. Can one define a natural larger class of maps, that would include F^{-1} in Example 8.1, for which Samuelson implies injective?

References

[1] S.S. Abhyankar, Polynomial maps and Zariski's main theorem, Oral presentation at the International Conference on Polynomial Automorphisms and Related Topics, Luminy, France, October 1992.

[2] K. Adjamagbo, H. Derksen and A. van den Essen, On polynomial maps in positive characteristic and the Jacobian conjecture, Proc. Amer. Math. Soc., to appear.

[3] H. Bass, F.H. Connell and D. Wright, The Jacobian conjecture: Reduction of degree and formal expansion of the inverse, Bull. Amer. Math. Soc. (N.S.) 7 (2) (1982) 287–330.

[4] H. Bass and G.H. Meisters, Polynomial flows in the plane, Adv. in Math. 55 (1985) 173–208.

[5] A. Bialynicki-Birula and M. Rosenlicht, Injective morphisms of real algebraic varieties, Proc. Amer. Math. Soc. 13 (1962) 200–203.

[6] J. Bochnak, M. Coste and M.-F. Roy, Géométrie Algébrique Réelle, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), Vol. 12 (Springer, New York, 1987).

[7] L.A. Campbell, Decomposing Samuelson maps, Linear Algebra Appl., to appear.

[8] L.M. Druzkowski, The Jacobian conjecture, Technical Report, Preprint 492, Institute of Mathematics, Polish Academy of Sciences, 1990.

[9] D. Gale and H. Nikaidô, The Jacobian matrix and global univalence of mappings, Math. Ann. 159 (1965) 81–93.

[10] K. Kurdyka and K. Rusek. Surjectivity of certain injective semialgebraic transformations of $\mathbb{R}^n$, Math. Z. 200 (1988) 141–148.

[11] G.H. Meisters, Jacobian problems in differential equations and algebraic geometry, Rocky Mountain J. Math. 12 (1982) 679–705.

[12] G.H. Meisters and C. Olech, A poly-flow formulation of the Jacobian conjecture, Bull. Polish Acad. Sci. Math 35 (11–12) (1987) 725–731.

[13] G.H. Meisters and C. Olech, Solution of the global asymptotic stability Jocobian conjecture for the polynomial case, in Analyse Mathematiques et Applications (Gouthiers-Villars, Paris, 1988) 373–381.

[14] G.H. Meisters and C. Olech, A Jacobian condition for injectivity of differentiable plane maps, Ann. Polon. Math. LI (1990) 249–254.
[15] H. Nikaidô, Relative shares and factor price equalization, J. Internat. Economics 2 (1972) 257–263.
[16] C. Olech, T. Parthasarathy and G. Ravindran, Almost N-matrices and linear complementarity, Linear Algebra Appl. 145 (1991) 107–125.
[17] T. Parthasarathy, On Global Univalence Theorems, Lecture Notes in Mathematics, Vol. 977 (Springer, New York, 1983).
[18] G. Ravindran, Global univalence and completely mixed glames, Ph.D. Thesis, Indian Statistical Institute, New Delhi, India 1986.
[19] K. Rusek, Polynomial automorphisms, Technical Report, Preprint 456, Institute of Mathematics, Polish Academy of Sciences, 1989.
[20] P.A. Samuelson, Prices of factors and goods in general equilibrium, Rev. Econom. Stud. 21 (1963) 1–20.
[21] R. Schramm, Boundary conditions for the univalence of NVL-transformations, Math. Ann. 236 (1978) 191–198.
[22] L. van den Dries, Alfred Tarski's elimination theory for real closed fields, J. Symbolic Logic 53 (1) (1988) 7–19.
[23] A. van den Essen, Polynomial maps and the Jacobian conjecture, in: CWI Tract., Vol. 84 (CWI, Amsterdam, 1991) 29–44.
[24] A. van den Essen and T. Parthasarathy, Polynomial maps and a conjecture of Samuelson, Linear Algebra Appl. 178 (1993) 191–196.